

PAPER

Multiple Sclerosis Impact Scale (MSIS-29): reliability and validity in hospital based samples

A Riaz, J C Hobart, D L Lamping, R Fitzpatrick, A J Thompson

J Neurol Neurosurg Psychiatry 2002;**73**:701–704



Additional information regarding this paper is available on the journal web site (www.jnnp.com)
See end of article for authors' affiliations

Correspondence to:
Dr J Hobart, Department of Clinical Neurosciences, Derriford Hospital, Plymouth, Devon PL8 8DH, UK; Jeremy.Hobart@phnt.swest.nhs.uk

Received 17 April 2002
In revised form 13 May 2002
Accepted 15 August 2002

Background and aim: The psychometric properties of rating scales are sample dependent and need evaluations in different samples. The Multiple Sclerosis Impact Scale (MSIS-29), a new patient based rating scale for multiple sclerosis (MS) was predominantly developed from a community based sample derived from the MS Society. A number of important patient characteristics of this sample remain unknown. The aim of the study was to evaluate five psychometric properties of the MSIS-29 in three hospital based samples: people admitted for rehabilitation, people admitted for intravenous corticosteroid treatment for MS relapses, and people with primary progressive MS.

Methods: People with MS were recruited from the three clinical settings. They completed several health measures. MSIS-29 data were evaluated for data quality, scaling assumptions, acceptability, reliability and validity, and compared with those from a previously reported community based study.

Results: A total of 233 people (rehabilitation =53; corticosteroids =76; primary progressive =104) completed questionnaires. In all samples, missing data were low ($\leq 2.2\%$), scaling assumptions were satisfied, and reliability was high (≥ 0.91). Correlations between the MSIS-29 and other scales were consistent with a priori hypotheses. Findings were consistent with those from the community samples.

Conclusions: The psychometric properties of the MSIS-29 are consistent across three hospital based samples, and similar to those in the community samples. These findings further support its use as an outcome measure in different clinical settings.

The Multiple Sclerosis Impact Scale (MSIS-29) is a new measure of the physical and psychological impact of MS from the patient's perspective (a copy of the MSIS-29 is on the journal web site; www.jnnp.com).¹ It was developed using the standard psychometric approach of reducing an item pool generated de novo from people with multiple sclerosis (MS). Psychometric evaluation of the scale was conducted in two large independent postal surveys of randomly selected, geographically stratified members of the Multiple Sclerosis Society.

In those samples, the MSIS-29 satisfied all recommended psychometric criteria for rigorous measurement.¹ However, the psychometric properties of health measurement instruments are sample dependent and cannot be established in a single study.² Moreover, the psychometric properties of the MSIS-29 were tested in randomly selected members of the MS Society. Although this sampling frame has the advantage of being truly representative of MS Society members, the disadvantage is that the percentage of people with a neurologist confirmed diagnosis of clinically definite MS cannot be confirmed. Furthermore, the disease type of those with MS and the representativeness of people who join charitable groups is unknown. This study aimed to examine the psychometric properties of the MSIS-29 in three hospital based samples of people with a confirmed diagnosis of MS in comparison with a community sample of people with MS.

METHOD

Sample

The sample was derived from adults with clinically definite MS who were consecutively admitted to the National Hospital for Neurology and Neurosurgery (NHNN) between February 2000 and July 2001 for rehabilitation and intravenous corticosteroid treatment of relapses, and a postal survey of people

from the Primary Progressive MS (PPMS) database of the NHNN. Patients in the rehabilitation and corticosteroid samples completed the MSIS-29 on admission along with other health measures and demographic questions.

People admitted for rehabilitation and corticosteroids were excluded from the study if they appeared to have severe cognitive impairment substantiated by neuropsychological testing, other comorbid disabling disorders, or were not English speaking.

Measures

All patients completed four self report measures: the MSIS-29 described above, the 59 item Functional Assessment of Multiple Sclerosis (FAMS³), the Medical Outcomes Study 36 item Short-Form Health Survey (SF-36⁴), and the 12 item General Health Questionnaire (GHQ-12⁵). In addition, rehabilitation and corticosteroid patients were rated by a neurologist on the Expanded Disability Status Scale (EDSS⁶).

The psychometric properties of the MSIS-29 were evaluated using standard methods. These methods are fully described in previous publications.^{1,7,8} In short, we examined five psychometric properties: data quality, scaling assumptions, acceptability, reliability, and validity.

RESULTS

Samples

Rehabilitation

Fifty eight people were consecutively admitted to the rehabilitation unit. Five people were excluded because of cognitive

Abbreviations: MS, multiple sclerosis; MSIS-29, Multiple Sclerosis Impact Scale

Table 1 Patient characteristics

Variable*	Hospital based samples			
	Rehabilitation	Corticosteroids	Primary progressive	Community sample
Number†	53	76	104	713
Sex				
Female	62	71	43	71
Age				
Mean (SD)	46 (12)	41 (11)	53 (10)	52 (12)
Range	23–69	17–83	28–73	18–82
Ethnicity				
White	94	87	95	98
Years since MS onset				
Mean (SD)	15 (10)	14 (11)	12 (6)	19 (11)
Range	0–35	1–60	3–32	1–59
Mobility indoors				
Walks unaided	4	53	19	32
Walks with an aid	49	43	54	40
Uses a wheelchair	47	4	28	28
Marital status				
Married	68	53	68	70
Living				
Living with others	83	82	85	81
Employment status				
Retired due to MS	55	26	63	56
Employed	21	55	15	19
Type of MS (%)				
Primary progressive	10	3	100	Unknown
Secondary progressive	31	17	0	Unknown
Relapsing-remitting	12	55	0	Unknown
EDSS				
Mean (SD)	7.1 (0.8)	5.6 (1.1)	Not available	Not available
Range	5.5–8.5	2.5–8.0	Not available	Not available

*All values are percentages unless specified otherwise. †For whom both physical and psychological scale scores could be computed.

impairment. A total of 53 people (91%) admitted for rehabilitation completed the questionnaires on admission.

Corticosteroids

Seventy eight people were consecutively admitted for intravenous corticosteroid treatment for relapses. Two people were excluded for cognitive impairment. A total of 76 (97%) people completed the questionnaires on admission.

PPMS

Questionnaires were sent to 119 people, 104 completed questionnaires were returned (87% response rate).

Table 1 shows the characteristics of the patients in the three samples. There were some differences in characteristics between the samples. The majority of patients in the rehabilitation and corticosteroid samples were women, whereas the majority of the PPMS sample was men. PPMS patients were older than the rehabilitation or corticosteroid samples, as might be anticipated.⁹ Other differences between the samples were consistent with expectations. The majority of patients in the rehabilitation sample used an aid or a wheelchair and the majority of patients in the corticosteroid sample walked unaided or used an aid. Many PPMS patients used an aid or a wheelchair, although some walked unaided. More than half of the corticosteroid sample was employed full time. More than half of the rehabilitation sample and an even greater proportion of the PPMS sample were retired because of their MS.

Data quality (table 2)

For all samples, missing data for items were low for both scales, and total scores could be computed for at least 99% of patients.

Scaling assumptions (table 2)

For all samples, frequency distributions for item response scales were quite symmetrical; items within each scale had

similar mean scores and standard deviations. Item-total correlations were satisfactory, although the lower range of item-total correlations of the physical scale in the rehabilitation and PPMS sample was <0.40. The physical scale of the rehabilitation sample had the lowest item-total correlations.

Acceptability (table 2)

For all samples, psychological scale scores spanned the entire scale range and scale scores were not notably skewed. Floor and ceiling effects were small and were considerably less than the recommended maximum of 15%.¹⁰

There were some differences between the samples for the physical scale scores. The lowest score was 3.8 (PPMS sample), 7.5 (corticosteroid sample), and 22.5 (rehabilitation sample). In this rehabilitation sample, only 75% of the MSIS-29 physical score range was used.

Although the mean scores were near the scale midpoint for all samples, the mean physical scale score of the rehabilitation sample was higher than the other two groups indicating greater physical impact. The standard deviation of the physical scale in the rehabilitation group was also lower than the other two groups. The mean psychological scale score of the PPMS sample was lower than the other two groups.

Reliability (table 2)

All estimates for Cronbach's α exceeded the standard criteria of 0.80.¹¹ The standard error of measurement is an estimate of error that can be used to calculate the 95% confidence interval (CI) for individual patient scores. For all samples, the 95% CIs for the MSIS physical scales were narrower than MSIS psychological scales, indicating that greater measurement precision is obtained when more items are used.

Validity (table 3)

Correlations between the physical and psychological scales of the MSIS-29 were 0.47 in the rehabilitation sample, 0.66 in

Table 2 Data quality, scaling assumptions, acceptability, and reliability of the MSIS-29

Psychometric property	Rehabilitation sample (n=53)		Corticosteroid sample (n=76)		PPMS sample (n=104)		Community sample (n=703)	
	MSIS-29 Physical	MSIS-29 Psychological	MSIS-29 Physical	MSIS-29 Psychological	MSIS-29 Physical	MSIS-29 Psychological	MSIS-29 Physical	MSIS-29 Psychological
Data quality								
Item missing data (%)	0–1.9	0	0–3.9	0	0–1.9	0–1.9	1.7–3.6	1.1–1.8
Computable scale scores (%)	100	100	100	100	99	99	98	99
Scaling assumptions								
Item mean scores	2.8–4.4	2.6–3.2	2.4–4.1	2.6–3.4	2.7–3.9	2.1–2.7	2.5–3.8	2.6–3.3
Item SD	0.9–1.5	1.3–1.5	1.1–1.5	1.3–1.5	1.0–1.5	1.1–1.3	1.2–1.6	1.3–1.4
Item-total correlation	0.26–0.64	0.49–0.74	0.46–0.79	0.54–0.88	0.33–0.81	0.54–0.84	0.60–0.86	0.49–0.77
Acceptability								
Scale range	0–100	0–100	0–100	0–100	0–100	0–100	0–100	0–100
Score range	22.5–97.5	0–100	7.5–100	0–100	3.8–100	0–100	0–100	0–100
Mean score (SD)	67.2 (17.4)	47.3 (24.3)	57.3 (23.1)	47.4 (26.1)	56.4 (23.3)	34.3 (24.5)	56.0 (26.6)	45.5 (25.2)
Mean score (SE)	2.4	3.3	2.7	3.0	2.3	2.4	1.0	1.0
Floor/ceiling effect (%)	0.0/0.0	1.9–1.9	0.0/1.3	2.6/1.3	1.0/0.0	1.9/2.9	0.9/3.9	1.7/1.9
Skewness	–0.525	–0.010	–0.41	–0.043	–0.17	0.87	–0.285	0.172
Reliability								
Cronbach's α	0.88	0.87	0.94	0.91	0.95	0.92	0.96	0.91
SEM*	6.0	8.8	5.7	7.8	5.2	6.9	5.3	7.6
95% CI†	±11.8	±17.2	±11.2	±15.3	±10.2	±13.5	±10.4	±14.9

*Standard error of measurement (SEM) = $SD \times \sqrt{1-\alpha}$. †95% Confidence intervals = $\pm 1.96 \times SEM$.

Table 3 Convergent and discriminant construct validity of the MSIS-29

		MSIS-29							
		Rehabilitation sample		Corticosteroid sample		PPMS sample		Community sample	
Instrument	Scale/dimension/variable	MSIS-29 Physical	MSIS-29 Psychological	MSIS-29 Physical	MSIS-29 Psychological	MSIS-29 Physical	MSIS-29 Psychological	MSIS-29 Physical	MSIS-29 Psychological
SF-36*	Physical functioning	–0.52	–0.28	–0.70	–0.44	–0.73	–0.30	–0.79	–0.41
	Mental health	–0.39	–0.73	–0.46	–0.77	–0.30	–0.64	–0.41	–0.76
FAMS†	Mobility	–0.63	–0.27	–0.75	–0.61	–0.75	–0.45	–0.88	–0.50
	Emotional wellbeing	–0.51	–0.67	–0.50	–0.74	–0.56	–0.75	–0.68	–0.68
GHQ‡	Total score	0.51	0.77	0.52	0.78	0.36	0.68	0.46	0.68
EDSS		0.27	0.14	0.69	0.48	–	–	–	–
Demographic variables	Age	–0.01	–0.06	–0.03	–0.16	–0.16	–0.09	0.22	0.03
	Sex	–0.01	–0.02	–0.20	–0.16	0.19	0.06	0.05	–0.05
	Years since diagnosis	0.09	0.09	0.18	–0.17	0.19	–0.02	0.19	0.03

*Medical Outcomes Study 36-item Short Form Health Survey: high scores = better health. †Functional Assessment of Multiple Sclerosis: high scores = better health. ‡General Health Questionnaire.

the corticosteroid sample, and 0.65 in the PPMS sample, indicating that the two scales measure related but distinct constructs.

Correlations with other measures and variables provide evidence for the convergent and discriminant validity of MSIS-29 as a measure of the physical and psychological impact of MS. The direction, magnitude and pattern of correlations were generally consistent with predictions. For example, in the corticosteroid and PPMS samples, the MSIS-29 physical scale correlated most highly with the FAMS mobility and SF-36 PF scales, and lowest with the SF-36 MH scale. In the corticosteroid sample, the MSIS-29 physical scale also correlated highly with the EDSS. In all samples, the MSIS-29 psychological scale correlated most highly with SF-36 mental health, FAMS emotional wellbeing and GHQ scales.

There were some notable differences in the pattern of correlations across the three samples. Correlations of the MSIS-29 physical scale with SF-36 PF, FAMS mobility, and EDSS were lower in the rehabilitation sample (0.27 to 0.63) than the other samples (0.69 to 0.88). Correlations of the

MSIS-29 psychological scale with FAMS mobility and EDSS were low as predicted in the rehabilitation and primary progressive samples (0.14 to 0.50), but higher in the corticosteroid sample (0.48 to 0.61).

DISCUSSION

The aim of this study was to evaluate the psychometric properties of the MSIS-29 in three hospital based samples, and compare these with the data from the community based sample. The psychometric properties of the MSIS-29 scales were similar in all four samples and demonstrate that the MSIS-29 satisfies standard criteria for reliable and valid measurement. The results demonstrate that the measurement properties of MSIS-29 in a range of hospital based samples are very similar to those obtained in a community context¹ and support the appropriateness of the instrument in a range of clinical settings.

The differences in mean scores among the three samples were as expected. Patients in the rehabilitation sample scored highest (worst) on the MSIS physical scale. Patients in the

PPMS sample scored lowest (best) on the MSIS psychological scale, supporting a previous finding that PPMS patients show better psychological functioning than patients with secondary progressive MS.¹² Vleugels *et al* suggested that this may be because PPMS patients feel they are more likely to know the rate of their physical deterioration, and that relapses are unlikely.

There were some other notable differences in the psychometric characteristic of the MSIS-29 among the samples. In the rehabilitation sample, the correlation between MSIS physical scale and the SF-36 PF scale and the EDSS was lower than expected. Closer inspection of findings in this sample revealed a substantial floor effect for the SF-36 PF (31%) and limited variability in EDSS score (range 5.5 to 8.5). In contrast, MSIS-29 physical scores spanned nearly the full scale range and showed minimal floor and ceiling effects. Even when cases with floor and ceiling effects were removed from the analysis, the correlation between the MSIS-29 physical and SF-36 PF was virtually unchanged at $r=-0.55$. This suggests that the floor and ceiling effects did not attenuate the correlations between scales. However, the limited score distribution of the SF-36 (range 0 to 65) may have attenuated the correlations between scales, as correlations are proportionally sensitive to scale ranges.

Although this study provides further evidence for the use of MSIS-29 as a rigorous measure of the physical and psychological impact of MS, ongoing evaluations are still required to further define its role in clinical practice and research. As traditional psychometric methods were used to develop and evaluate the MSIS-29, and it is difficult to interpret small differences in performance across samples, it is important that newer psychometric methods such as Rasch item analysis¹³ and Item Response Theory models¹⁴ are applied to the MSIS-29 data. These will generate item calibrations in the different samples permitting a more extensive comparison of scale performance.

ACKNOWLEDGEMENTS

We wish to thank the people with MS who freely participated in the studies that provided the data used in this study, the Multiple Sclerosis Society of Great Britain and Northern Ireland for their collaboration, and Gordon Ingle for help with the PPMS database. This

study was funded by a grant from the NHS Health Technology Assessment Programme (but the views and opinions expressed do not necessarily reflect those of the NHS Executive) and the MS Society of Great Britain and Northern Ireland.

Authors' affiliations

A Riazi, J C Hobart, A J Thompson, Neurological Outcome Measures Unit, Institute of Neurology, London, UK

J C Hobart, Department of Clinical Neurosciences, Derriford Hospital, Devon, UK

D L Lamping, Health Services Research Unit, London School of Hygiene and Tropical Medicine, UK

R Fitzpatrick, Department of Public Health and Primary Care, University of Oxford, UK

Competing interests: none declared.

REFERENCES

- 1 **Hobart JC**, Lamping DL, Fitzpatrick R, *et al*. The Multiple Sclerosis Impact Scale (MSIS-29); a new patient-based outcome measure. *Brain* 2001;**124**:962–73.
- 2 **Stewart AL Hays RD**, Ware JE Jr. The MOS Short-Form General Health Survey: reliability and validity in a patient population. *Med Care* 1988;**26**:724–35.
- 3 **Cella DF**, Dineen K, Arnason B, *et al*. Validation of the functional assessment of multiple sclerosis quality of life instrument. *Neurology* 1996;**47**:129–39.
- 4 **Ware JE Jr**, Snow KK, Kosinski M, *et al*. *SF-36 Health Survey manual and interpretation guide*. Boston: Nimrod Press, 1993.
- 5 **Goldberg DP**, Hillier VF. A scaled version of the General Health Questionnaire. *Psychol Med* 1979;**9**:139–45.
- 6 **Kurtzke JF**. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;**33**:1444–52.
- 7 **Hobart JC**, Lamping DL, Freeman JA, *et al*. Evidence-based measurement—Which disability scale for neurologic rehabilitation? *Neurology* 2001;**57**:639–44.
- 8 **Hobart J**, Freeman J, Thompson A, Kurtzke scales revisited: the application of psychometric methods to clinical intuition. *Brain* 2000;**123**:1027–40.
- 9 **Thompson AJ**, Polman CH, Miller DH, *et al*. Primary progressive multiple sclerosis. *Brain* 1997;**120**:1085–96.
- 10 **McHorney CA**, Ware JE Jr, Lu JF, *et al*. The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;**32**:40–66.
- 11 **Nunnally JC**, Bernstein IH. *Psychometric theory*. 3rd edn. New York: McGraw-Hill, 1994.
- 12 **Vleugels L**, Pfenning L, Pouwer F, *et al*. Psychological functioning in primary progressive versus secondary progressive multiple sclerosis. *Br J Med Psychol* 1998;**71**:99–106.
- 13 **Rasch G**. *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press, 1960.
- 14 **Lord FM**, Novick MR. *Statistical theories of mental test scores*. Reading, MA: Addison Wesley, 1968.